DOCUMENT RESUME

ED 464 366 CS 511 126

AUTHOR Moscovitch, Edward
TITLE Evaluation of the Alabama Reading Initiative. Final Report.
INSTITUTION Alabama State Dept. of Education, Montgomery.
PUB DATE 2001-12-00
NOTE 55p.; Prepared by Cape Ann Economics.
AVAILABLE FROM Alabama Reading Initiative, Alabama Department of Education, PO Box 302101, Montgomery, AL 36130-2101. Tel: 334-353-1389; Web site: http://www.alsde.edu.
PUB TYPE Reports - Research (143)
EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Administrator Role; Elementary Secondary Education; Evaluation Methods; Literacy; Professional Development; Program Effectiveness; Reading Achievement; *Reading Improvement; *Reading Instruction; *Reading Programs
IDENTIFIERS Alabama Reading Initiative

ABSTRACT

A study evaluated the third year (2000-2001) of the Alabama Reading Initiative (ARI), which is designed to achieve 100% literacy by targeting reading performance of beginning reading and first-grade students, to expand the reading power for second- through twelfth-grade students and to intervene for struggling readers at all grade levels. Reading test scores on the Stanford Achievement Test, Level 9 were reviewed over a span of 5 years. The evaluation was also based on personal visits to 15 ARI schools and 10 after-school focus groups, which included some 225 teachers from 45 schools around the state. Results indicated: schools participating in the ARI outperformed schools not in the program; 14 of the schools in the first year of the program lost ground in the third year, and the second year gains of the 59 schools that started the program in the second year were smaller than their first year gains, suggesting annual gains level off over time; the top schools outperformed the norm by 4 NCE (normal curve equivalency) points the first year and 2.8 points the second, while students in the lowest 5 schools fell 1.7 NCE points the first and 1.6 the next year; and teachers reported enthusiastically that their faculty was working together as never before. Focus group discussions indicated that successful schools have in common: a strongly committed principal; a full-time reading specialist; faculty members who work cooperatively; teachers who make frequent use of diagnostic tests; teachers spend the greater part of the day in small-group instruction to meet individual needs; the school teaches phonemic awareness and systematic phonics to all students in the primary grades and to struggling readers in all grades; teachers re-enforce comprehension skills for all students; the school actively encourages students to read; the school receives follow-up and guidance from the state ARI staff; and the system's central office is supportive. Findings suggest that the second year of the program was the most successful; a better measure of student progress is needed than the Stanford 9 test; and the ARI will have to invest in a large corps of state-funded reading specialists who can keep in constant touch with new ARI schools or those that are struggling. An appendix contains a more detailed quantitative analysis. (Contains 5 references and 13 notes.) (RS)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 464 366

Evaluation of the Alabama Reading Initiative

Final Report

By Edward Moscovitch

December, 2001

T.E. Inovar, Tr.

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it.
- Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

CS 511 126

CAPE ANN ECONOMICS

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

Table of Contents

Introduction *1*
Plan of the Report *1*

SUMMARY OF FINDINGS AND RECOMMENDATIONS **2**

FINDINGS – QUANTITATIVE DATA 2
Interpreting the Results *4*
Variation in Scores Across Schools *4*
Results by Grade Level *6*
Conclusions from the Quantitative Data *7*
FINDINGS – QUALITATIVE DATA – SCHOOL VISITS AND FOCUS GROUPS 8
Whole-School Improvement *8*
Expecting More of Children *9*
Elements of Success *10*
RECOMMENDATIONS 11
The Crossroads – Invest in Success, or Program Dissipation *11*
Summary of Recommendations *12*
The ARI Diagnostic Test *12*
Cost of Recommendations *13*
Is It Worth It? *14*

WHOLE SCHOOL REFORM **15**

THE ARI – AS TEACHERS SEE IT 15
Love of Reading *15*
Teacher Growth, Enthusiasm, and Co-operation *16*
Balanced Approach to Reading Instruction *16*
SUCCESS FACTORS 17
The Principal *18*
The Reading Specialist *20*
Faculty Teamwork *21*
Individual Assessment *22*
Individual and Small-Group Instruction *22*
Teaching Phonics and Phonemic Awareness *23*
Encouraging Students to Read *24*
Reading across the Curriculum *25*
Follow-up from the ARI *25*
Central Office Support *25*

RECOMMENDATIONS FOR BROADER SUCCESS **25**

Maintain Two-Week Training, but Improve Speakers *27*
Full-Time Reading Specialists in Every School *27*
Certification for ARI Reading Specialists *28*
ARI Instruction and Teacher Certification *29*
ARI Renewal Training Every 3 Years *29*
More Funding For Follow-Up Training And School Visits *30*

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

Mandatory, diagnostic, computerized tests 31
Gold Star Schools 33
Principal Training 34
Central Administrator Training and Commitment 34
Summer Reading Institutes 35
Reading Curriculum Materials and Software 35
More ARI Staff 36

APPENDIX 1 – QUANTITATIVE ANALYSIS IN MORE DETAIL **36**

USING STANDARDIZED TEST SCORES TO EVALUATE INDIVIDUAL SCHOOLS 36
Results from the "Value-Added" Approach 38
More Detailed Analysis 44

CAPE ANN ECONOMICS

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

Introduction

As outside evaluator for the Alabama Reading Initiative (ARI) in its third year, I have reviewed Stanford 9 reading test scores for half a million Alabama students for the last 5 years. This evaluation is also based on personal visits to 15 ARI schools and 10 after-school focus groups, which included some 225 teachers from 45 schools around the state.

Although the schools visited represent only a small fraction of the 424 schools participating in the ARI, these schools and those invited to the focus groups were broadly representative, selected to include some of the highest and lowest performing ARI schools, as measured by the Stanford 9 reading scores. School visits have been made in every corner of the state, from Arab in the Northeast to Prichard in the southwest. They include rural schools and urban; all-black schools, integrated schools, and all-white schools.

It is impossible not to notice – and be impressed with – the sense of excitement and accomplishment reflected by the great majority of teachers I've met. The ARI is not just about reading; it's become a highly effective movement for whole school revitalization. Scores of teachers have shared their pride in what they've accomplished – kindergarten classes where essentially everyone reads by the end of the year, schools where library circulation has doubled or more, behavior problems and special education referrals reduced substantially and interest in reading up substantially. Most teachers report that the students entering their classes are noticeably better prepared than they had been before the reading initiative. Teachers tell me that they have a better sense of how the various elements of a reading program (phonemic awareness, phonics, fluency, comprehension) fit together, how to assess children's individual needs, and how to devise strategies to meet those needs. They report that their faculties are working more co-operatively than had previously been the case. Their students are doing more independent work; as teachers they've been pleasantly surprised at how much their children are capable of. Some of the schools I've visited rank with the best and most vibrant schools I've seen – anywhere.

Plan of the Report

Those pressed for time will get the major ideas of the evaluation in the Summary of Findings and Recommendations, beginning on the next page. Later sections of the report spell out in more detail the difference the initiative has made in teaching and student reading, the common elements that characterize the most successful ARI schools, and the report recommendations.

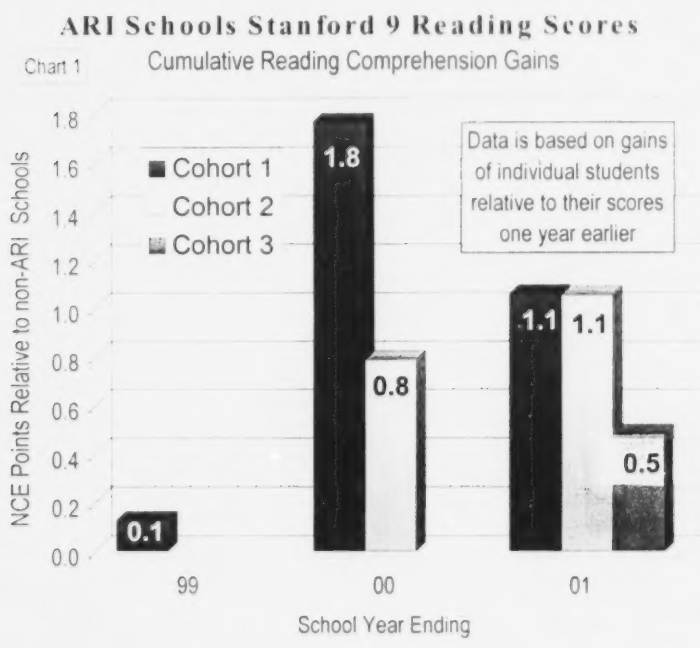
CAPE ANN ECONOMICS

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

SUMMARY OF FINDINGS AND RECOMMENDATIONS

Findings – Quantitative Data

The only statewide data available on student reading scores is from the Stanford 9 tests, which are given in every Alabama school from grades 3 through 11. No reasonable inference on individual school performance can be drawn simply from the absolute level of student scores in any school, since there is such a strong correlation between the socio-economic status of school students and their test scores. A better measure – the one used in this evaluation - can be taken by following the performance of individual students from one year to the next and comparing schools based on how much progress their students made.



By this measure, schools participating in the ARI outperformed schools not in the program, as shown in Chart 1 at left. The chart is based on the normal curve equivalents reported for the Stanford 9. In the 98-99 school year, there were 14 ARI schools whose students took the Stanford 9, shown in dark blue on chart 1 (these are referred to as cohort 1 schools).[1] In the spring of 1999, students in grades 4 through 11 in these schools had shown a gain over the prior year 0.1 NCE points greater than students in non-participating schools.[2]

[1] There are 16 schools in Cohort 1, but 2 are primary schools with no students in grade 3 or above.

[2] For example, if the average Alabama 3rd grader scored 50 NCE points (the median of the national distribution) and the average Alabama 4th grader also scored 50 NCE points, an ARI 4th grader who scored 50 in 1998 would score 50.1 in 1999 while her counterpart in a non-participating school would score 50. Because the tests begin in grade 3, we cannot measure gains until grade 4.

BEST COPY AVAILABLE

CAPE ANN ECONOMICS

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

In the second year, students in the 14 cohort 1 had gained an additional 1.7 NCE points relative to students in non-participating schools. The chart shows cumulative gains; by the second year, students in these 14 schools were scoring 1.8 NCE points (which translates to about 3.5 percentiles for students in the middle of the distribution) higher than students elsewhere in the state who started with similar scores.

In preparing this chart, student gains are first calculated separately at each grade level. ARI student gains are compared to gains in non-ARI schools grade by grade, and then the results are combined across grades.[3]

In the 1999-2000 school year, 59 additional schools with students in grades 4 and higher joined the ARI. Students in these schools showed performance gains in this first year 0.8 NCE points greater than students in non-participating schools, as shown in yellow on the chart.

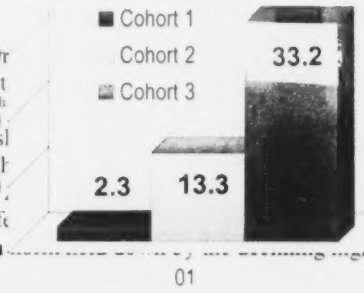
In the 2000-2001 school year, the students in the cohort 1 schools lost ground relative to students in non-participating schools (for convenience, the performance of students in non-participating schools will be called "the norm."). In that one year, their relative performance slipped by 0.7 NCE points. Taken over 3 years, however, students in the cohort 1 schools were 1.1 NCE points above the norm (which translates into about 2 percentiles).

Students in the cohort 2 schools made further gains in their second year; they gained an additional 0.3 NCE points and were 1.1 NCE points above the norm by the end of their second year.

In the 2000-2001 school year, an additional 163 schools with students in grades 4 and above joined the ARI. Their students outperformed those in non-participating schools by 0.5 NCE points.

ARI School Enrollment

Chart 2 Thousands of Students, 2000-2001



The ARI is a very large program, as shown in Chart 2. By the third year of the program, there were 2,270 students in cohort 1 schools

[3] Scores for ... since average scores of Alabama students on the Stanford 9 test ... vary from grade to grade. Third graders score right at 50, but 4th ... Thus, the average Alabama 4th grader (not in ARI schools) sl... er a loss of 1.5. The gains of ARI schools are compared to th... students in the high school grades score in the mid 40s. Since the ... ts, any comparison that failed to take these grade to grade diff... aggerate the ARI performance, since it would be compared to ... school scores.

BEST COPY AVAILABLE

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

taking the Stanford 9 test who had also taken it the year before. There were over 13,000 such students in cohort 2 schools and over 33,000 students in cohort 3 schools.

In no sense, then, is the ARI simply a pilot program. Rather, it is a massive effort to change reading instruction for tens of thousands of Alabama students.

Interpreting the Results

Unfortunately, the Stanford 9 results do not tell us the absolute level of a child's performance (for example, that a 4th grade student is reading at a 5th grade level). We can only infer that a student who scores at the same percentile in 4th grade as in 3rd grade has made normal progress for the year. Averaging across all 3 cohorts, ARI students outgain students in non-participating schools by 0.5 NCE points per year. If these results were to continue over the course of a 12-year school career, a student would show a relative gain of 6 NCE points (which is equal to about 11.5 percentiles for an average student). Thus, if students in a non-participating school are roughly at the national average throughout their career, a student in an ARI schools who started at the 50th percentile might finish his career at the 62nd percentile.

From 3 years worth of data, we cannot know whether average gains of 0.5 NCE points per year can be maintained. The fact that the 14 cohort 1 schools lost ground in 2001 and that the 2nd year gains of the 59 cohort 2 schools were smaller than the 1st year gains suggests that the annual gains at any one school will level off over time. This is hardly surprising or even disappointing; if a school takes a struggling reader 2 grades below level and brings him up to grade level and then maintains him on grade for several years, it has done well.

Variation in Scores Across Schools

Simply comparing average performance of ARI schools with the performance of non-ARI schools masks the enormous variation in how well ARI schools are doing. The scores of the very best schools in each cohort are compared with the lowest scoring schools in chart 2 below

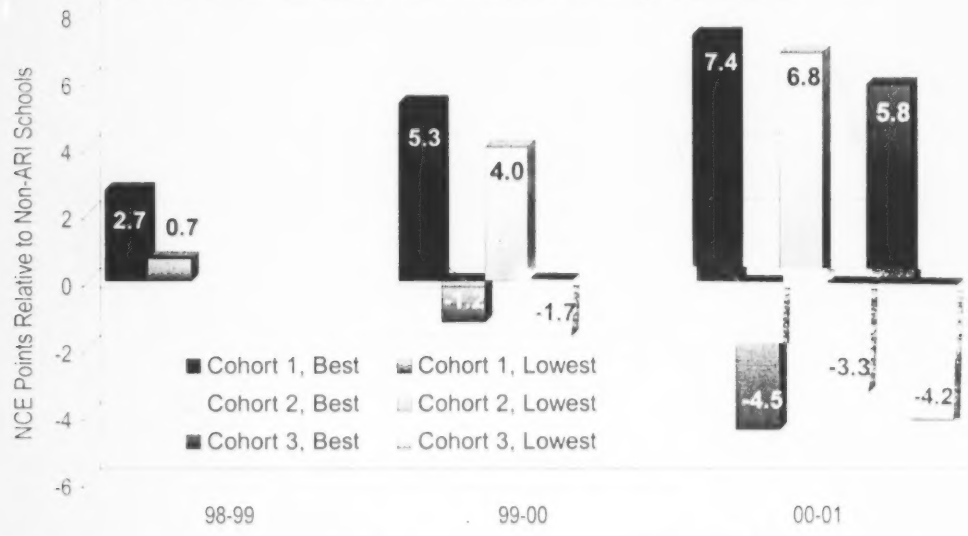
CAPE ANN ECONOMICS

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

Variability in ARI Schools Stanford 9 Reading Scores

Chart 2

Best and Lowest Schools by Cohort, Cumulative Reading Comprehension Gains



For Cohort 1, the chart shows the average of the 3 top scoring schools and the 3 lowest schools. As in chart 1, this chart measures the gains in scores of students who took the test for 2 consecutive years, and compares ARI schools with non-ARI schools. Thus, the top 3 Cohort 1 schools raised their student scores by 2.7 NCE points (about 5 percentiles) more than the norm in the first year, outgained non-participating schools by another 2.6 points the second year, and moved ahead by yet another 2.1 points the third year. By the third year, these students were 7.4 NCE points ahead of comparable students in non-ARI schools; had the students started at the 50th percentile they'd now be performing at the 63rd percentile.

The lowest Cohort 1 schools gained 0.7 NCE points the first year, but lost this and more the 2nd year (falling 1.9 points so they were 1.2 points below the norm). They lost still more ground the 3rd year, finishing 4.5 NCE points below the norm.

Since there are 59 Cohort 2 schools reporting test results, I've averaged the top 5 and lowest 5 schools. The top schools outperformed the norm by 4 NCE points in their first year and by an additional 2.8 the second year. Students at the lowest 5 actually fell 1.7 NCE points behind students at non-participating schools; they fell another 1.6 points the next year. We have only one year of data for Cohort 3 schools; the 5 best showed a relative gain of 5.8 points; the 5 lowest a relative loss of 4.2 points.

BEST COPY AVAILABLE

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

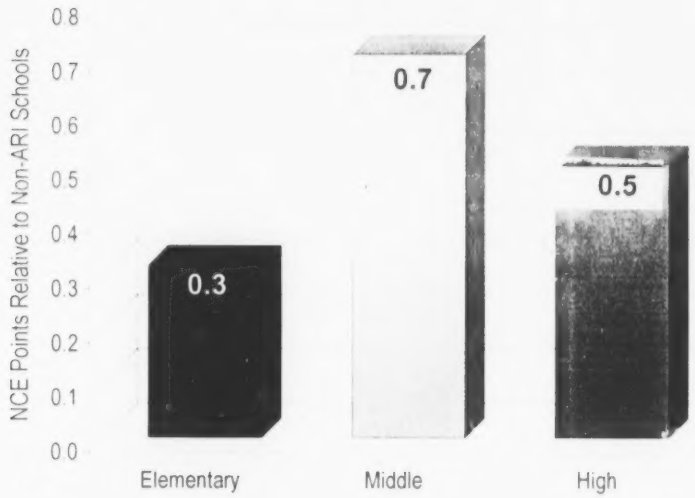
The lesson is clear – some ARI schools are doing very well indeed; others are not.

Results by Grade Level

Most, but by no means all, ARI students are in elementary schools. Of ARI students tested by the Stanford 9 in the spring of 2001, 26,000 were in elementary grades, 13,600 in middle grades, and 9,000 in high school.[4]

ARI Schools Stanford 9 Reading Scores

Chart 3 Average Reading Comprehension Gains, by Grade Level



As shown in Chart 3, the average annual reading gain (relative to non-participating schools) is actually highest for the ARI middle schools at .7 NCE points gain per year. The high schools are next, with a .5 point relative gain. The ARI elementary schools outperform non-ARI schools, but by a smaller margin. Since gain is measured relative to prior year performance, and since the Stanford 9 isn't given until 3rd grade, we can measure gain only for students in 4th grade and above. For a K-5 elementary school, this means we are measuring the relative increases for only 4th and 5th graders. We know nothing about how well a school is doing with the roughly 2/3 of its students who are in Kindergarten or grades 1 through 3, nor do we have any data about ARI primary schools.

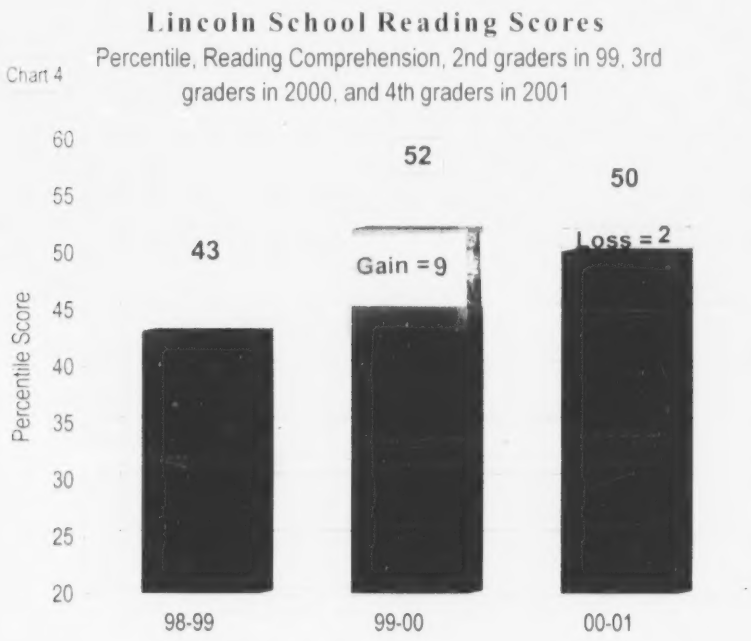
This partial measure may be unfair to schools that do a particularly good job of raising scores of students in the early grades.

An example of this is provided by following the students who were in 2nd grade at the Lincoln Elementary School in Talladega in the 1998-1999 school year. The school itself paid to

[4] These scores are for students who've taken the test two years in a row, so exclude grades 1 through 3. The total number of elementary students in ARI schools is therefore well in excess of 50,000.

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

have the Stanford 9 administered in 2nd grade; the class tested at the 43rd percentile in reading comprehension, as shown in Chart 4.



The next year the Lincoln school joined the ARI; as 3rd graders this class averaged 52 in reading comprehension, for a gain of 9 percentiles. This gain was invisible to the statewide testing process, since it does not test in 2nd grade. As 4th graders, this class fell slightly, to the 50th percentile. At the time, the Lincoln School did not have 5th grade. This drop in the scores from 3rd to 4th grade was therefore the only data on individual student gains available for the school, and it was negative. However, the school's overall performance with this cohort from 2nd grade to 4th grade was quite impressive. In a way, the school was actually penalized for how well it improved its students in 3rd grade, since there was less room to improve further in 4th grade.

At the very least, the example of the Lincoln School suggests the limitation of a program that does not begin until 3rd grade. This drawback is particularly important for a literacy program that understandably puts a great deal of its energy into students in the primary grades.

Conclusions from the Quantitative Data

On average, ARI schools outperform schools not in the ARI. Many ARI schools are performing very well and some are truly outstanding. On the other hand, some ARI schools are doing no better than (and in some cases not as well as) schools not in the ARI. Given this finding, the school visits and focus group interviews focused on the characteristics that successful schools have in common and also the characteristics of unsuccessful schools, to identify the key factors in determining school success.

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

The Stanford 9, the only available quantitative measure of ARI school success, begins in 3rd grade and therefore it tells us nothing about the success (or failure) of teaching in the early grades. This is one of the major factors behind the key recommendation, discussed with other recommendations below, that the ARI develop an evaluating tool designed to measure in all grades reading achievement in all grades.

The data imply that 1999-2000 was the most successful year for the ARI, which suggests that more needs to be done to make sure the program is not watered down as it expands.

Findings – Qualitative Data – School Visits and Focus Groups

The evaluation included 15 school visits. In each case I walked through the school with the principal, visiting classes and talking with teachers, and spent an hour or more talking with a small focus group of teachers. These interviews were held without principals or reading specialists in attendance, and teachers were very helpful and candid. In addition, there were 10 after-school focus groups, each with 20 to 30 teachers from 4 or 5 ARI schools in the area (other than the ones visited personally). These, too, were held without administrators or anyone from the ARI central office, and were also very candid, free-wheeling sessions.

Whole-School Improvement

Although the ARI is nominally about reading, it is really a movement that aims at whole-school revitalization, centered on literacy. In countless schools around the state, the combination of whole-faculty training in the summer followed by solid leadership from the principal, active help for teachers from a knowledgeable reading specialist, frequent assessment of individual student needs, and small-group instruction designed to meet individual needs has brought the faculty together and made a major impact not only on student reading but on teaching methods and student achievement across the board.

Teachers reported enthusiastically that their faculty is working together as never before. The cumulative impact of such excitement from so many teachers over the course of four weeks is compelling. The following quote sums it up:[5]

> *"I love the reading initiative; I think its great. We're closer as a faculty and go to each other to talk about what we're doing."*

The literature on school improvement describes how the traditional notion - the teacher closes her classroom door and is an island unto herself, while with the principal is a building

[5] A later section of the report explores these changes in more detail, using a number of quotes to let the teachers express their enthusiasm for the ARI in their own words.

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

manager who puts out fires but leaves teachers alone – is one of the major obstacles to school improvement.[6] Part of the genius of the ARI is that it has gotten past teacher isolation to get faculties working in harmony. The two weeks of training in the summer, with ample time for teachers to spend relaxed time together, is a critical part of this. Another important element is that the ARI isn't just another state program, it's a *movement*. Teachers, principals, and reading specialists in the more successful schools feel they're part of something broader than their school alone. At some schools, 5 or more teachers have trained to be ARI instructors and spend part of their summers teaching faculties of new ARI schools. Teachers visit other ARI schools and reading specialists in each part of the state meet regularly.

The quantitative data summarized above suggest that the ARI was most successful in the 1999-2000 school year and didn't do quite as well in 2000-2001. There's a ready explanation. In 1999-2000, every school had a state-funded, full-time reading specialist. Specialists were not included in the program's first year; by the third year there wasn't sufficient funding to provide one for every school. At the same time, the program was still relatively small in its second year, so that ARI schools had frequent contact with the ARI director (whose enthusiasm and knowledge have played a major part in the program's success) and with the ARI central staff. The central staff has not expanded apace as the program has grown. It added 267 schools in 2000-2001 and 423 schools in 2001-2002. As a result, many schools feel isolated and suffer from lack of follow-up, frequent contact, and guidance from the ARI staff.

To be successful as a larger program, the ARI will have to invest in a large corps of state-funded reading specialists who can keep in constant touch with new ARI schools or those that are struggling, and also to find ways to keep ARI faculties and principals in frequent contact with each other. In this way, the ARI can preserve itself as an active movement of enthusiastic teachers.

Expecting More of Children

The initiative has made a major difference for students.

> *"We are making massive changes in the way we teach reading – and massive changes in the lives of our students....We're creating lifetime readers."*

I saw this personally at several schools. Golden Springs Elementary School covers one of the lower-income areas of Anniston, itself a low-income city. The town is hard-pressed financially, so the school has no art, music, or gym teacher. Still, I ate in the lunchroom (along with a 100 or more students) and was able to talk with the principal in a normal tone of voice.

---

[6] See, for example, "Building a New Structure for School Leadership", by Richard F. Elmore, published in 2000 by the Albert Shanker Institute

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

As the students (mostly low-income, mostly black) finished eating, they quietly took out their books and started to read quietly.

Those who work with the ARI on a daily basis should not take such scenes for granted. When this story is told outside Alabama; the response is always one of astonishment and admiration.

As students accept more responsibility and do more work independently, their teachers' sense of what is possible expands. Rising student performance also means increased teacher enthusiasm.

Elements of Success

The focus groups confirm what the quantitative results tell us – many schools succeed, but some do not. Successful schools have in common most if not all of the following elements:

- The principal is strongly committed to the reading initiative and knows how to provide educational leadership in the school,
- The school has a full-time reading specialist with in-depth, hands-on reading instruction experience,
- Faculty members work co-operatively instead of competitively. Faculty members at each grade level meet frequently together (at least weekly) and meet periodically with teachers in the next lowest and next highest grades,
- Teachers make frequent use of diagnostic tests to identify needs of individual students,
- Teachers spend the greater part of the day in small-group instruction to meet individual needs,
- The school teaches phonemic awareness and systematic phonics, to all students in the lower primary grades and to struggling readers in all grades,
- Teachers re-enforce comprehension skills for all students, not only in the language block or in language classes, but throughout the school day and across the entire curriculum. This means that art, music, and gym teachers, as well as middle and high school history, math, science, and social studies teachers are also teaching reading skills,

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

- The school actively encourages students to read, using a wide variety of appealing books,
- The school receives follow-up and guidance from the state ARI staff,
- The system's central office is supportive.

Teachers and principals report that they come back from the initial ARI training overwhelmed with how much there is to do. Frequent follow-up from the ARI staff and frequent contact with other ARI schools is a key factor in school success.

Recommendations

The Crossroads – Invest in Success, or Program Dissipation

The ARI is at a crossroads. The fundamental concept is sound and scores of outstanding schools around the state can point to it as a key factor in their success. However, it has expanded faster than its funding and infrastructure; as a result many participating schools are adrift. They don't have a full-time reading specialist, they haven't received follow-up and assistance from the ARI, support from the principal or district central office is inconsistent, they aren't doing frequent diagnosis and don't know how to do small-group instruction. Not surprisingly, there Stanford 9 reading scores are disappointing.

If the ARI continues to expand without adequate resources, it will have more and more schools that are indistinguishable from their non-ARI neighbors and fewer and fewer additional faculties will see any reason to undergo the summer training in return for minimial financial support and a reduced likelihood of success.

Given the ARI's enormous potential, it would be a shame to see this happen. As one teacher put it,

> *"This is such a good idea; I hate to see it fade. The Department of Education shouldn't just give ARI training to schools and walk away."*

Instead, Alabama should invest in success. The evidence suggests that the great majority of schools can succeed if they receive adequate funding, training, and support. As this happens, more and more faculties across the state will want to join the movement, and the ARI will be able to insist that principals and central offices demonstrate a commitment to take the steps necessary for success.

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

Summary of Recommendations

Specifically, the ARI should:

- Provide state funding for a full-time reading specialist in every ARI school,
- Set up a certification and training program for ARI reading specialists to assure that all state-funded specialists have in-depth, hands-on experience in the theory and practice of research-based reading instruction,
- Increase per-school funding to $5,000 per year to cover more teacher release time for visiting other ARI schools and participating in follow-up professional development during the school year,
- Provide more frequent outside help to ARI schools by increasing the ARI central staff to include 9 reading professionals and maintaining at least one regional reading specialist for every 20 participating schools,
- Require every participating school to renew its ARI membership every 3 years, including attendance of at least 85% of the faculty and the principal at one-week "refresher" summer institutes,
- Establish 20 ARI Gold Star Schools, to be selected competitively. These outstanding schools would receive additional grants of $100,000 per year to hire one extra reading specialist or assistant principal and to cover additional curriculum materials or faculty release time. In return they would serve as sites for principal and reading specialist training,
- Require specific training for principals and central office administration as a condition for ARI participation and funding; this training would include several days visiting ARI Gold Star Schools, and
- Develop a test and diagnostic vehicle to be administered by all ARI schools.

The ARI Diagnostic Test

The overwhelmingly enthusiastic response from ARI teachers (including many teachers at schools that show relatively low student gains in relation to the norm) suggests that more is going on than the Stanford 9 is measuring. This underscores the importance of a test that is designed specifically to measure what the ARI is trying to achieve and that can be administered to students at all grade levels.

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

This new test should serve both to tell the ARI's story and also to implement the ARI central tenet that frequent diagnosis of individual student needs is essential to educational success. Once they know what individual students need and have the practical know-how to meet those needs, the great majority of teachers will work hard to meet those needs.

The ARI should develop a test with the following characteristics:

1. It is diagnostic, indicating major reading needs of individual students,

2. It is vertically scaled, so that it can measure student reading levels at least from grade 1 through grade 8,

3. It can be administered to each individual student with an average teacher time commitment of 15 to 20 minutes, and

4. It is completely computerized, so that results are instantly available, for individual students and by class, to the teacher, the principal, and the ARI itself.

Each ARI school should be required to administer this test to each student at least 3 times a year. A computerized test needn't be a simple multiple-choice test; modern software allows the kind of notations teachers now make with paper and pencil in a running record or the QRI (now used by many ARI schools) to be made with the computer keyboard instead, with the results instantly available for analysis.

For the great majority of ARI schools that already are doing regular assessment, this new test can replace the existing test, so it does not become a new requirement. Because the new test will be so closely involved with instruction, the ARI staff should seek input from a broad range of teachers and reading specialists from ARI schools in putting together this new assessment vehicle.

For schools that don't now do regular assessments, the new test will increase substantially the odds on success.

Cost of Recommendations

Applied to the existing base of ARI schools, implementation of these recommendations will cost an additional $23 million per year, raising total program costs to $34 million annually. For each 100 schools added to the program, spending should be increased by an additional $7 million per year.

Part of this additional cost could be obtained by redirecting some of the current spending of the Department of Education. For example, participating schools that are eligible for Title 1

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

funding under the ESEA (a program which is explicitly about literacy) might be required to use at least ¼ of their ESEA funds to support a reading specialist and other ARI costs. This makes educational as well as economic sense, since teachers have made clear that reform works best when a school adopts a single focus and sticks with it for several years. Literacy is obviously central to education; until a school excels at teaching reading, it is unlikely to succeed in other areas. Thus, it makes sense for the department to concentrate its attention and funding on successful implementation of the reading initiative. It makes even less sense to have competing reading initiatives. Teachers in some counties have reported that county reading programs are at variance with the ARI. Spending funds on two competing (and possibly incompatible) programs reduces the likelihood that either will succeed.

Is It Worth It?

This report is written at a particularly difficult time for the Alabama state budget, so $23 million seems like a lot of money. However, it really is a small amount compared to total state and local education spending in the state – or even to total federal assistance to Alabama education. The track records of the best ARI schools shows that the program has a dramatic potential to revitalize schools and create a generation of Alabama readers. The funding required is a small price to pay for this kind of success.

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

WHOLE SCHOOL REFORM

This evaluation is based on discussions with some 350 teachers from across Alabama. The great majority of them are very enthusiastic about the reading initiative and the difference it has made to their schools, their students, and their teaching. The best way to convey their enthusiasm is to let the teachers speak for themselves:

The ARI – As Teachers See It

The following are as close as possible to exact quotes, excerpted from the interview notes taken during school visits and after-school focus groups. The quotes are typical of the sentiments expressed by the teachers at these groups. As the teachers themselves make clear, morale is high in most ARI schools – children are reading as never before and teachers are working more effectively and co-operatively.

Love of Reading

"We are making massive changes in the way we teach reading – and massive changes in the lives of our students....We're creating lifetime readers."

From a teacher who's been at it 18 years – "This is the best thing I've seen – it warms my heart to see lagging readers want to read in science."

"Never before in my life have I seen students so motivated to read."

"Now when I come in in the morning, my kids are sitting on the floor reading....and during the day they beg me to let them read."

Book circulation at the Ruhama Elementary School library went from 1600 books a year before the ARI to 13,000. Teachers from several schools reported library circulation doubling or tripling. Many schools set goals – often as high as 110 books per student per year.

"I don't have any children who don't love to read."

"Now the children are so focused and involved they don't have time to get into trouble." Teachers from several schools report dramatic declines in discipline referrals. At West Blocton Elementary, for example, referrals are down 70%.

Teacher Growth, Enthusiasm, and Co-operation

"I learned more in the two weeks of ARI training about teaching children to read than I learned in four years of college." "The reading initiative is the best professional development in all my years of teaching." "I knew many of the individual elements of a good reading program, but I didn't know how they fit together."

One teacher feels like a new teacher; she postponed her retirement, even though her salary is less than a thousand dollars more than her retirement benefit.

"It's been tremendous to see how the faculty has come together."

"I was a good teacher before, but now I'm a better teacher, and it's because of the ARI training."

"Many of us were so tired of giving tests and seeing zeroes – we wanted something that would work. I saw immediate results – it was fun to teach."

"The Reading Initiative makes me want to be a better teacher."

"I didn't learn how to manage three groups at one once (to do small group instruction) in college."

"I love the reading initiative; I think its great. We're closer as a faculty and go to each other to talk about what we're doing."

"When you walk in the door, it's a different school (compared to before the reading initiative)." "Our faculty is more on track than we've ever been. It brought us together as a faculty – we all understand what each is talking about. We are excited about teaching; we want to go to staff development."

"Our school is more interactive than before; teachers are giving the QRI and targeting individual needs. There are more co-operative groups and more writing; students are immersed in print."

"Now I know how to meet the needs of those behind in reading." "We love the reading initiative; we'd do it again!"

Balanced Approach to Reading Instruction

"We picked books that were pretty, but didn't teach students to read....I really see the importance of a strong phonics background."

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

"I always taught phonics, and thought this was the starting point. Now I realize that phonemic awareness has to come first, and I see a big difference in my students."

"The college where I got my degree was whole-language oriented, and they gave us the sense we should avoid systematic phonics and phonemic awareness."

A kindergarten teacher reports she's changed her whole approach to teaching. In particular, she's focused more on phonemic awareness and has more tools. She was a whole-language teacher and thought she was doing phonics. She now realizes she wasn't doing enough to meet students' needs. Understanding sounds is more important than knowing letters.

Success Factors

There's little mystery why some schools do well and others don't. The schools that do well share all (or most) of the following characteristics:

- The principal is strongly committed to the reading initiative and knows how to provide educational leadership in the school,
- The school has a full-time reading specialist with in-depth, hands-on reading instruction experience.
- Faculty members work co-operatively instead of competitively,
- Teachers monitor the progress of struggling readers with periodic diagnostic tests that identify needs of individual students,
- Teachers find ways to provide small-group instruction to meet individual needs,
- The school teaches phonemic awareness and systematic phonics, to all students in the lower primary grades and to struggling readers in all grades,
- Teachers re-enforce comprehension skills for all students, not only in the language block or in language classes, but throughout the school day and across the entire curriculum. This means that art, music, and gym teachers, as well as middle and high school history, math, science, and social studies teachers are also teaching reading skills,
- The school actively encourages students to read,

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

- The school receives follow-up and guidance from the state ARI staff,

- The system's central office is supportive of the reading initiative and actively involved in its implementation.

The Principal

Both from my personal observation and from the comments of the many teachers I've met, it's clear that the principal plays a critical role in the success of the best ARI schools. This is hardly surprising in view of everything we know about the importance of principals.

Some specifics:

- The principal must make the reading initiative her top priority and must stick with it for several years. If this is simply the fad of the month, it won't work. At two schools visited, the principal made clear that nothing – not suspensions, errands, sending papers to the office, or field trips – could interrupt the morning reading block.

- The principal makes sure that teachers do frequent assessment of individual student needs and is actively involved in using the data to monitor individual student progress.

- The principal must hold his teachers accountable. The teachers want him to do this – they want him to make sure that everyone in the school is with the program. We forget that failure by the 2nd grade teachers will affect the ability of 3rd grade teachers to do their job. Some principals do this in part by reviewing weekly lesson plans or holding regular meetings with grade-level teams. Others require teachers to report on who in their class needs individual attention, what their needs are, what the teachers are doing to meet them, and how well it's working.

- While holding teachers accountable, the principal must give her teachers latitude to implement the reading initiative in the way they think best.

- The principal has to be a visible presence in his school. In most, but not all, of the successful schools, the principal visits every classroom several times a week and, if at all possible, at least once a day. At most of the unsuccessful schools, teachers report that they are not likely to see the principal in their class more than 2 or 3 times a year.

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

- The principals at the best ARI schools have a passion for helping children learn to read. These principals typically give their teachers freedom to decide how to implement the ARI program, but teachers understand that the school's underlying commitment to better reading is not open to debate.

- The more personal the principal's support the better. For example, a few principals take one of their school's reading groups. Most have paid considerable attention to the research on how children learn to read and discuss it with their staffs. Some use "reading to the principal" or "lunch with the principal" as a reward for good student work.

- The principal should lead by example. One principal brings his book to a different class each morning to read quietly during the school's 15 minute silent reading period. Another handles the morning assembly herself, leading the children in songs, teaching them sign language, and reading to them. Students come to class calmed down and ready to learn.

- The principal's passion for reading should include a knowledge of the subject. Teachers respect and admire a principal who is up on the professional literature and is constantly sharing ideas with them. One principal leads a regular faculty study group on reading instruction.

- Teachers need praise, recognition, and encouragement. A good principal has to be a cheerleader for his school. One principal raises money in the community and uses some of it to thank her teachers by taking them to dinner.

- A successful principal goes to the community to raise extra funds, so he can provide teachers what they need to succeed.

Why are teachers willing to follow really good principals? Here are some answers from teachers who admire their principals:

- The principal works harder than we do.

- She trusts us and gives us a real say in running the school.

- She treats us in a fair and professional manner.

- He's always there to give me what I really need.

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

The principal must somehow give reasonable flexibility while still insisting on the most important changes. One teacher explained that she had resisted changes in teaching techniques that her colleagues were making and that the principal preferred. He gave her flexibility to continue on her own for a few months, but ultimately came to her and, in the nicest possible way, made it clear that she'd have to join the team.

The Reading Specialist

Teachers report they come back from the initial summer training enthusiastic but overwhelmed. There is unanimous agreement that the process works best if they have a full-time reading specialist in their school. This reading specialist can be the intellectual spark-plug of the changeover. It falls to her to help teachers figure out new strategies when the first things they try don't work. She should spend a good part of her day modeling for teachers – giving them hands-on ideas about how to get their message across.

In a school where things are working well, classroom teachers will talk with the reading specialist two or three times a week, asking for help in dealing with a particular student or for ideas on how to put across a certain lesson. Often, the classroom teachers will set their own agenda for the kind of help they need from the specialist.

The most effective specialists will often help the principal devise major changes in the way the school day is structured.

A really good specialist will take the lead in assuring that individual student assessment is done regularly and done well. She will work with teachers to analyze assessment data and devise strategies for meeting individual student needs. The reading specialist can play a pivotal role in teaching her colleagues what small-group instruction is and how to do it.

It is possible, but far more difficult, to make the initiative work without a full-time specialist. Generally, where there is no specialist, the faculty makes an effort to figure out how to apply what they've learned on their own. Sometimes they succeed, sometimes not.

Ideally, the reading specialist should spend some of her time working with the most challenged readers in small groups. This is particularly true for struggling readers in the upper grades, where 2 or 3 students may need the most basic instruction, including phonemic awareness and systematic phonics that would no longer be needed by the great majority of their classmates.

An important part of the reading specialist's role is to bring in a continuous stream of new ideas to the school faculty. This should include coming up with techniques to use with particular children when the classroom teacher has tried everything she knows and is at a dead end. Some specialists maintain an informal "library" of professional books and articles that

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

her colleagues can borrow. In successful schools, the reading specialist regularly shares with the faculty what she's learned at the regular monthly sessions held by the ARI.

The reading specialist has to be an integral part of the school team, available on a casual basis to talk with teachers several times a week. It's hard to get this kind of communication if the specialist has to divide her time among 2 or 3 (or in some cases 4 or more) separate schools.

In some cases, schools have simply promoted one of their existing faculty to fill a vacancy in the reading specialist position, even if that person has had no reading instruction beyond the 2-week ARI training. Teachers want someone with advanced training and – most important – with hands-on reading experience. Naturally, they want someone who believes in the ARI principles they are trying to implement.

Faculty Teamwork

Teacher teamwork is a key factor in successful ARI schools. If each teacher closes her door and does her own thing, the initiative is almost certain to fail. It can only work when all the teachers in a particular grade talk to each other frequently and co-ordinate their teaching. Innovation is difficult, but it goes better when several people work on it together. This means looking to each other for suggestions and encouragement, and sharing successful strategies. Cohesive teams at each grade level can provide help to new teachers in the school and make sure that they are part of the reading effort. Absent this kind of unity, the enthusiasm (and knowledge) the faculty obtains at the ARI training will dissipate in a few years as trained teachers leave and are replaced by new teachers who may go to ARI training, but weren't part of the original bonding experience.

A central part of this co-operation is regular meetings of the grade-level teams. In many of the best schools, this is done with common planning time, preferably several times a week. At some schools, the principal requires team meetings on a weekly basis and teachers meet informally several times a week.

Schools work even better when teachers in one grade regularly talk with the teachers in the grades above and below about what each grade team needs to do to prepare students for the next level. It seems an obvious point, but if the teachers in, say, second grade all have slightly different curricula and learning goals, the third grade teachers are likely to lose several weeks in the fall bringing all their students to the same starting point. With regular meetings like this, the pressure for each team to work together comes not only from the principal, but also from faculty peers in higher grades.

At one unsuccessful ARI school, teachers remain competitive with one another and deliberately refuse to share information because they want students in their class to score higher

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

than those in the class next door. Overcoming this kind of selfishness is a major challenge for good principals.

Individual Assessment

One of the key elements of the ARI is that each teacher should have quick ways of assessing the reading progress of each student, with particular emphasis on the skill deficiencies of struggling readers. These assessments should not only give the child's reading level but also let the teacher know exactly where the child's problem lies. The cycle of frequent assessment, tailored instruction, and re-assessment to determine progress is a very powerful idea. It lets the data – not ideology – determine the kind of instruction each child receives. If a reliable diagnostic test says that Mary is not a fluent reader (all too frequent, in middle school as well as primary grades), then she needs instruction and practice specifically designed to increase fluency.

Diagnostic tests are also important because they help build faculty support for the ARI and all the hard work and change successful implementation requires.

> *"Once we understood how far behind many of our children were and what they needed, of course we wanted to do what it takes to help them to read."*

Diagnostic tests can also be a part of a principal's accountability program. Not only do the teachers know what each students needs; they also know that the principal knows, and that he'll be looking at future assessments to make sure students are making progress.

All too frequently, teachers in higher grades fail to diagnose the reading difficulties of their students. One 4th grade teacher explained that until the reading initiative she and her colleagues hadn't understood that many of their students were way behind in reading. In any case, it hadn't been her job to teach reading, since at that grade teachers concentrate on content areas, like science, math, and social studies. Another teacher pointed out that older children develop coping strategies to disguise their reading difficulties.

Individual and Small-Group Instruction

Small group and individual instruction geared to the specific needs of struggling readers is a central part of the ARI philosophy.

One approach is for teachers to learn to do small-group instruction. At several schools visited for this evaluation, teachers spend over half the school day working with their class divided into small groups. Take Brewton Elementary School's first grade as an example: a quarter of the students will be working on reading skills with the teacher (almost certainly she'll have grouped them by reading level). Another quarter will be reading books of their

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

choice independently; a third quarter will be doing workbook exercises focusing, for example, on identifying words with short vowels (the day I was there the exercise was to color in those boxes -and only those boxes - that had a picture of something using short "o"). The fourth group will be using phonics practice software program (that looked to them like a computer game).

In upper grades, one group of students might be be using the Accelerated Reading program of independent reading and computer-based comprehension tests, a second group might be working together to answer questions from the teachers on the characters and plot of a book she's assigned, a third might be having a student-led literacy circle, while the 4th will be reading aloud and working on fluency and comprehension with the teacher.

Teachers who've tried small group instruction feel very strongly about it:

> *"Now that I see how this is done, I'd never teach any other way." "I know exactly where each of my students is and what kind of help they need from me."*

A group of teachers can take collective responsibility for students during the language block, with students grouped by reading levels. For example, a school with 4 5th grades might have two teachers working with the more advanced 5th grade readers in relatively large classes. Meanwhile, the other two teachers (and possibly the reading specialist) would be working with smaller groups of struggling readers. In a more sophisticated implementation of this, one of the smaller groups might be for students with difficulty in decoding (who need phonics and phonemic awareness) while the other might be for students who need help with comprehension (for example, asking students why they are reading a given piece and what they hope to learn from it).

West Blocton Elementary School uses the computer lab to facilitate small-group instruction. Each grade team has its own lab and each teacher gets the lab for about an hour and a half each day. This time is divided into thirds; for 30 minutes a small group of students works with the teacher while the other 2/3 of the class is getting computerized practice in reading.

Schools with a reading specialist may have the specialist take students who are particularly far behind and "pull" them to offer highly specialized, small-group instruction. While clearly helpful, this small-group instruction is a complement to small group instruction by the classroom teacher rather than a substitute for it.

Teaching Phonics and Phonemic Awareness

Successful schools are teaching phonemic awareness in kindergarten and first grade. In some schools, teachers introduce a sequential series of books with decodable text, often starting in

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

the second half of the kindergarten year. By the end of kindergarten, most students who follow this approach are reading.

All too often, teachers in grades 3 and beyond, when asked about struggling readers, report that the difficulties are comprehension-related. Of course, many students in these grades do need help in comprehension. But most schools – even ARI schools – do not consider the possibility of phonics help, even when the students are painfully slow in decoding. In one school that had seen a drop in scores, teachers reported that a particular struggling reader could decode, so they were concentrating on comprehension strategies. But she could read aloud only haltingly, at half the normal speed, and clearly needed help with fluency. The school had no diagnostic instrument in place to pinpoint her areas of need. By way of contrast, the most successful schools had programs to provide systematic phonics and phonemic awareness to struggling readers in upper grade. Some had special materials to offer first grade phonemic awareness concepts with examples and pictures that would be more comfortable to children in 4th, 5th, or higher grades.

Encouraging Students to Read

Good schools find ways to encourage children to read. Many use Accelerated Reader, and report that children like the program and enjoy having a choice of books. Many offer prizes to students who read a large number of books successfully. These prizes can be as simple as a shiny pencil or a gold star; they might include reading to the principal, having lunch with a favorite teacher, permission to eat candy in class, or even a class pizza party.

Some schools encourage students to have a book with them at all times. The Arab Junior High School requires students to have at least one book checked out of the school library and to take a book home with them at night. The Abrams School in Bessemer requires students to take a book home, to read to their parents, and to bring back a parent's signature affirming that this had been done. The Hemphill in Birmingham sends contracts home to parents, requiring them to read to their children for seven minutes a day at 7 pm seven days a week (the "triple seven" program).

In younger grades, the best schools make sure that children are read to daily. The Leon Sheffield in Decatur has a retired gentleman who comes in daily to read to children and to work with struggling readers. A kindergarten teacher at the Woodmeade School in Decatur called members of her church until she had five volunteers, one for each day of the week, to come read to her children for an hour. For older readers, volunteers can agree to be read to, helping students with fluency and comprehension.

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

Reading across the Curriculum

Teachers can re-enforce reading skills while teaching history, science, and mathematics. For example, teachers in every subject can be required to use the "word of the week" in their lessons. Teachers in almost any subject can have students read aloud and can re-enforce the skills needed for good comprehension.

At the Somerville Road School in Decatur, the gym teacher has a word wall. He's painted the letters of the alphabet on the gym floor in three different sets, with individual letters spaced the width of the gym. Instead of capture the flag, students play games that involve running to the letters of a particular word.

Follow-up from the ARI

As indicated above, teachers come back from the ARI training overwhelmed with the magnitude of the changes that need to be made. They want frequent advice and help from the ARI staff. Schools in the first cohort reported the excitement of that initial year and the help (and morale boost) they got from ARI Director Katherine Mitchell's frequent visits. One principal from a school with disappointing scores (from Cohort 3) reported that he had asked several times for help, but that, as of October of his second year, he still hadn't had a visit from anyone from the ARI.

Central Office Support

Consistent support from the district central office can be an important factor in a school's success. Absent such support, the central office can change principals just after a school has gone through training, thereby destroying leadership continuity at the school, particularly if the reading initiative is not a high priority of the new principal.

In Mobile County, the school central office started its own reading program that, according to teachers, is somewhat at variance with ARI training. It isn't helpful if the District's outside lecturers tell teachers to ignore what they learned a few months earlier at ARI training.

Teachers in some districts thought the central office could help by trying to minimize the paperwork burden placed on principals and encouraging them to spend more time in classes.

RECOMMENDATIONS FOR BROADER SUCCESS

The Alabama Reading Initiative is at a cross-roads. Its basic idea is sound – two weeks of training in research-based reading instruction is a powerful first step. Many faculties are able

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

to take the ideas they learned at training and run with them. The ARI can point to some spectacularly successful schools.

But other faculties need more. For want of leadership from the principal, support from a good reading specialist, failure to use diagnostic testing, not moving toward small group of instruction, lack of follow-up from the ARI staff, or failure to teach an appropriate mix of phonics and comprehension, they fail to make gains.

In effect, Alabama is following a low-budget, low-commitment strategy. It provides good (very good) initial training. But it has only 1 ARI central office professional staff member for every 85 participating schools, along with one regional in-service center specialist working with (but not hired by) the ARI for every 40 schools.

The state funds an average of 1 reading specialist for every 5 schools. It seeks no serious commitment from central offices or principals. Some superintendents change principals with abandon. There isn't enough prior training for principals or central office administrators.

In the long run, this is not a viable strategy. As ARI school scores fall to the state average, fewer and fewer schools will find the ARI attractive and the legislature will eventually lose interest.

Letting the ARI fade away would be a terrible mistake. Properly funded and supported, the ARI has demonstrated that it can be a smashing success. The experience of scores of schools and the testimony of hundreds of teachers affirm this.

The Department should instead take the high-investment, high-commitment, high pay-off approach. It should fund full-time reading specialists for all ARI schools. It should increase the annual grant to give schools more flexibility to arrange follow-up support and professional development. In return, it should set tough standards for reading specialists, require principal training, and establish a mandatory diagnostic test that lets teachers know what their students need and also gives the Department the data it needs to tell the ARI story.

This high-investment, high-commitment approach should lead to a high success rate among ARI schools. With the higher grants and the higher success rate, schools would be clamoring to join the program. With increased demand, the department would be in a position to require the appropriate level of commitment.

The changes recommended here would increase Alabama's costs by roughly $23 million a year with the current number of schools, or $30 million with 100 additional schools. This will seem like a lot of money, particularly during the current recession and state fiscal crisis. But in the grand scheme of things, this is small money indeed – even for a poor state like

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

Alabama – compared to the enormous payoff from lifting reading performance and creating a generation of Alabama readers.

Maintain Two-Week Training, but Improve Speakers

A few of the teachers interviewed said that the 2-week summer training was too long. But since so very many teachers report that this training has made such a big difference in their school – not only in terms of teaching teachers how to teach reading, but also in terms of faculty bonding and morale – there seems little reason to change the broad outline of the training program.

Teachers who went to training in the first year clearly liked being "off-campus" together for several days. Serious consideration should be given to holding training sessions, at least for one of the two weeks, at some college campus or even some (modestly priced) resort. The ability to relax together after class, to eat dinner together, and to "hang out" in the evening are all very positive to faculties, who don't normally get to do this sort of thing. If this approach is taken, it might be possible to shorten the training by one day, since there'd be more time together in the evening.

Many teachers reported that the quality of instruction was uneven. They'd like ARI to make sure that presenters have hands-on classroom experience and are able to present "how-to" examples about specific things they can do to meet their students' needs. One useful suggestion was that ARI staff should visit the classrooms of possible presenters to see them in action.

Two weeks out of their summer is a big commitment for teachers to make, particularly in light of the low pay they receive. In return for such a commitment on the teachers' part, ARI should invest more staff time in making sure that the presenters are as good as they can be.

Full-Time Reading Specialists in Every School

Through the ARI office, the state of Alabama should support a full-time reading specialist in every participating school. A reading specialist who is part of the faculty "family" at the school is a critically important factor in school success. Almost all of the really successful schools visited had a full-time specialist; lack of a full-time person was always a factor at schools that have not been successful (this was true at schools visited personally; it was also the first factor mentioned at focus groups by teachers from less successful schools).

One reason a full-time specialist is important is simple time – between working individually or in small groups with students who are particularly far behind in reading, modeling for teachers, analyzing student diagnostic tests, working with teachers to build programs to meet

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

individual student needs, and keeping up with the professional literature and with the successes and failures of near-by ARI schools, a teacher needs full time in any one school.

But there's more at work here than time alone. As one reading specialist who covers more than one school explained, she's only at one of her two schools 2 or 3 days a week, and then only for 2 or 3 hours. As a result, she doesn't really feel a part of that school's faculty and isn't comfortable recommending the kinds of changes she thinks the school needs. The same can be true for teachers. Going to someone else to say that you can't meet a particular student's situation and that you need help isn't easy; you have to have a close and trusting relationship of a kind that's difficult to build with someone who isn't regularly in school. Also, teachers want to be able to have frequent, informal talks with the reading specialist. The teachers are themselves tied up in class most of the day; it works best if they know they're likely to run into the reading specialist in the halls or after school. If the specialist is only there one day a week and the teachers need to make a formal appointment, much of the spontaneity and frequent response that are so valuable will be lost.

The Somerville Road School in Decatur has two full-time reading specialists (they got the second one by "trading in" a reading recovery teacher). These two seem to gather extra excitement and momentum from each other. In larger, high-poverty schools such as this one (the school has 500 students and an 80% poverty rate), the ARI should offer a small bonus to encourage the school to transform an existing position into a second specialist.

There's another powerful advantage to implementing this recommendation. With this change, the value of designation as an ARI school will be enhanced considerably. More schools will want to join the program. This will allow the ARI to be more demanding in what it requires of member schools. The combination of more reading specialists and higher standards for admission to the program will increase the odds on success. And, with the program made more attractive both because of the higher grants and higher success rates, the ARI will still have plenty of new schools to choose from.

Certification for ARI Reading Specialists

As matters now stand, many of the reading "specialists" in ARI schools are not really reading specialists at all, but teachers (often good ones) who've been promoted from the ranks and who've had no special training in reading other than the two-week ARI summer course.

The reading specialist is not just another warm body. Her effectiveness depends on her knowledge and her ability to help students and teachers. Hiring reading specialists who weren't properly prepared was a major issue raised by the teachers interviewed.

ARI should require that specialists have extensive classroom experience and be thoroughly versed in the research and best practice on reading instruction.

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

Specifically, the Department of Education should develop the ARI Reading Certificate. To get it, a teacher would have to have a certain amount of classroom experience (five years at a minimum), have college-level (and graduate-level) courses in reading instruction, consistent with the NICHD research, and have gone through a few weeks of training with ARI director Katherine Mitchell and other top ARI instructors.

This certification would not necessarily be the same as receiving a graduate degree in reading from an Alabama university. However, Auburn, the University of Alabama, and other universities would be encouraged to modify their curriculum and course content to satisfy ARI requirements. One problem here is that actual course content might not match the stated curriculum. To deal with this, there should be some sort of in-service training included in the certification requirement and some kind of practical evaluation of candidates' knowledge.

Some sort of interim requirement should be available for current ARI reading specialists, who might be given two or three years to meet the requirements.

Implementation of this recommendation would have a major impact on the teaching of reading in Alabama, since it would be a powerful incentive for most (hopefully all) education schools in the state to modify their curriculum and tie it into the ARI.

ARI Instruction and Teacher Certification

Several teachers in the focus groups suggested that no one should be able to receive an Alabama teaching certificate without attending an ARI 2-week summer training session – an excellent suggestion.

ARI Renewal Training Every 3 Years

The bonding effect and morale boost from ARI training is highest in the year immediately following training. Over time, these impacts fade. Inevitably, teachers leave and are replaced by new teachers. Generally, these new teachers do attend training, but not with colleagues from their own school. All too often, principals leave and are replaced by principals who weren't part of ARI training. There's no guarantee that the ARI and reading instruction will be a priority for the new principal.

For these reasons, participation in the ARI should be granted for a 3-year interval only; to "renew" in the program a school's faculty and principal should be required to attend renewal training every 3rd summer. In this way the faculty then in place would have a common bonding and training experience. This training needn't be for a full two weeks; there's no need to teach over again what the faculty has already learned. A week-long program would be suffi-

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

cient or, even better, a 4-day residential program starting at, say, noon on Tuesday and lasting until mid-afternoon Friday.

This idea has been "test marketed" with scores of teachers in the focus groups and has received an enthusiastic response. Despite the inconvenience of leaving home, most teachers prefer the residential setting because they appreciate the value of evenings together.

Teachers say they'd welcome the chance to get more training, and to receive it as a faculty. Since most have already had the basic training, it would make sense to offer a program in which about half the courses were electives, so teachers could choose workshops on whatever subjects they felt they needed help with. The remainder of the program could cover major new developments in the field and allow time for common planning with the principal.

More Funding For Follow-Up Training And School Visits

When they come back from the 2-week ARI training, teachers (and principals) are overwhelmed with the magnitude of the transformation they've been asked to make. They all report that they've wanted more follow-up help.

There are two separate but inter-related recommendations here:

1. Schools need, and want, more visits from ARI staff. This is helpful at all times, but critically important in the first few months. Indeed, more than one teacher suggested that at least some of these visits be unannounced so that ARI staff would know what's really going on. This will require additional ARI staff, discussed below.

2. Teachers would like more follow-up training and more opportunities to visit each other's classes and classrooms in nearby ARI schools to get ideas. One suggestion I particularly liked was to let faculties bring back the instructor or instructors of their choice from summer training for additional sessions during the year. The current funding allowance to each school for ongoing training should be increased from the current $1,800 to $5,000 a year, with the understanding that much of this could be used for substitutes to allow for cross-school visits.

It is easy to forget how overwhelming the reading initiative's program for reading instruction and curriculum is. Teachers come back with a 3-inch notebook crammed with ideas on how to teach reading. They are expected to monitor student progress and to assess the individual needs of struggling readers. For many, perhaps most, of the teachers, phonemic awareness is a whole new idea – even for primary grade teachers. Certainly for teachers in higher grades the idea that they are responsible for instruction in reading is new. For any teacher, giving individual or small-group instruction to students who are 2, 3, or more years behind in reading is daunting.

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

One group of teachers (from a school making no progress) told me that they know there is a better way to teach reading and that it will produce better readers, but they're still not sure exactly what it is and they have no idea how to apply it in their own classrooms. This school needs (and hadn't yet received) several days of on-site help with specific, hands-on techniques teachers can use. Teachers and principals alike made clear that at this early stage they needed not only the help of their own reading specialist but outside help and a strong assist from the ARI staff itself.

Mandatory, diagnostic, computerized tests

The ARI should require that all participating schools administer 3 times a year a computerized, diagnostic test. This test should have the following properties:

- It should be diagnostic and indicate individual student needs and shortcomings in all areas of reading, including phonemic awareness, phonics, and reading comprehension.

- It should give student reading levels, from non-readers through 8th grade. In this way, it should allow the ARI to track student progress from one year to the next.

- It should take no more than 20 minutes per student of the time of a teacher, teacher aide, or reading specialist (if some of the test were multiple-choice or other computer problems and some involved reading to an adult, the total time for the student could exceed 20 minutes).

- It should be completely computerized, with results instantly available to the teacher but with the ability to present whole-class summaries to the teacher, comparisons across classes to the principal, and comparisons across schools to the district central office and the ARI. It should build a central, pupil-specific database that allowed the schools and the ARI to track the progress of individual students over several years.

In terms of what it can accomplish in relation to its cost, this is the single most important recommendation in my report.

The idea of doing frequent diagnosis of individual student needs and teaching accordingly is one of the most powerful ideas in the ARI summer training. Without exception, the successful schools I visited were giving diagnostic tests and teachers were using what these tests told them to structure individualized instruction.

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

Use of these tests underscores the message from ARI training that each child's needs are unique. A properly designed test that covers both phonemic awareness and phonics on the one hand and comprehension and fluency on the other gets beyond the ideology of the reading wars. If Mary has trouble with phonemic awareness, then she needs phonemic awareness instruction, regardless of what the more extreme whole language partisans might say. And if Robert doesn't comprehend well, he needs help with comprehension strategies, regardless of how important phonics may be. The great majority of teachers want to do what's best for their students; if a diagnostic test tells them that Sarah needs help with reading fluency, the teacher will find a way to give her that help.

Even in schools where teachers don't hold ideological positions on reading instruction, there's a tendency not to recognize decoding difficulties in older students and to think that struggling readers in the higher grades need only comprehension help; a good diagnostic test will resolve this in a nice way.

These diagnostic assessments will also help with accountability. The teacher will have an immediate picture of what each child in her class needs. She'll also know that the principal knows as well. And she'll know that another test will be given in a couple of months, and that the principal will be able to see which students have made progress and which haven't. For that matter, the principal will know that the district central office and the ARI will know whether his school is making progress.

Because the test is diagnostic, the accountability benefits aren't simply that a teacher or a principal knows she can be held accountable. The principal and reading specialist will be able to find common patterns of success and failure. For example, a particular second grade teacher might have generally good results, with most of her students making good progress. But if the test shows that the 2 or 3 who are having most trouble with comprehension aren't moving ahead, this will let everyone know that the teacher needs help in this area. Or perhaps the tests will show that what look like comprehension problems for most struggling readers in grades 4, 5, and 6 have their roots in lack of fluency and a poor background in systematic phonics, suggesting that the reading specialist and upper grade teachers need to do more work in this area.

Finally, this recommendation is important because the ARI doesn't have a good way of telling its story in a quantitative way. Teachers generally feel the program is a success and report that they can see their students reading far better than before the ARI. But the Stanford 9 results don't really tell this story. ARI school scores are higher than non-participating schools, but the difference in scores is not as large as one would expect after visiting ARI schools and talking with enthusiastic ARI teachers. Perhaps the ARI techniques make teachers feel better about themselves but haven't made a major dent in reading problems. More likely, students really are reading better but not in ways that yet show up on 3rd and 4th grade

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

tests. Either way, given the investments being made in the ARI – as much in teacher time and effort as in state funding – and the overwhelmingly positive reports of teachers in ARI schools, the ARI needs a measure of its success (or failure) that's tied directly to what ARI schools are trying to accomplish.

Instead of adding yet another layer of testing, this recommendation allows the same test to do triple duty – to help the teacher know her class, to help the principal monitor her teacher's success, and to allow the ARI to tell its story.

This idea has been "test-marketed" with at least 100 teachers at several of the after-school focus groups and received a favorable response. We can therefore be confident that the idea will be well received, particularly if a group of teachers and reading specialists from ARI schools have a chance to participate in the selection and design of the test vehicle.

The ARI should write up a set of specifications as to what it wants its test to accomplish and then put the project out to bid. The ARI currently includes 424 schools. This is a big enough market to get the attention of test vendors. There's no doubt that what's suggested here is technically feasible. The Lexia comprehensive reading test, currently in use in some Alabama schools, is 95% of the way there, including both reading substance and computer software. The DIBELS test also used in many Alabama schools meets most of the reading requirements and is strongly grounded in the NICHD research, but is not available with computerized scoring. Perhaps with a market this large the DIBELS founders would be willing to develop or lease the appropriate software. The QRI is used by the greatest number of ARI schools, particularly for older students. In its current form it takes about 45 minutes to an hour; with this sort of incentive, its designers might find a way to use computerization to reduce the adult time needed for each student to the desired 20 minutes. The STAR is already computerized but is not diagnostic; if its sponsors were interested they might find a way to form a consortium with other tests that meet the diagnostic requirements.

Purchased for over 400 schools, these existing tests (modified to meet the ARI specifications) might be obtained quite reasonably – Lexia's licensing fee is only $5,000 per school the first year and less than $1,000 for annual maintenance.

Gold Star Schools

The best ARI schools are simply wonderful. Visiting these schools would surely be a source of inspiration to principals and reading specialists across the state. However, it would be unfair – and counter-productive - to choke their success by drowning them in visitors.

ARI should establish 20 or so "ARI Gold-Star Schools." Schools selected for this honor would receive an extra $100,000 per year, with the understanding that the better part of this would cover hiring an additional senior faculty member (either an assistant principal or a

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

second reading specialist, at the school's discretion). The remainder could be used at the school's discretion, for books, materials, teacher release time, professional development, or any other educational purpose.

In return, gold star schools would agree to be sites for principal and reading specialist training.

The ARI would hold a competition for gold star schools, with some kind of judging panel to make the selection.

Principal Training

Teachers at the focus groups for less than successful schools report that their principals rarely visits classes and in many cases don't place high priority on the reading initiative.

To increase the odds on success, the ARI should insist that principals at schools accepted into the ARI agree to a two to three-week training program. This would be in addition to the 2-week training for the school faculty, which the principals should also be required to attend in its entirety with their faculties.

At least one week of this principal's training should be spent visiting ARI Gold Star schools. Each school visit should be for the entire day, with ample time for the visiting principals to talk with the school principal and reading specialist, but also with time to visit classes, talk with teachers, and observe reading instruction.

Other topics of training should include a review of the NICHD research on how children learn to read, principals as education leaders, and how successful principals win the respect of their faculties. Principals should be exposed to small group instruction and how to help teachers move in this direction. Finally, it would make sense to give the principals themselves hands-on experience in some basics of reading instruction so that they can handle small intervention groups for struggling readers and otherwise lead by example when they return to their schools.

Central Administrator Training and Commitment

Given the importance of consistent support from the District central office, the superintendent (in smaller districts) or the deputy superintendent responsible for curriculum and professional development should be required to attend a training program that included visits to some of the Gold Star schools.

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

The ARI should make clear to school central offices that, barring an emergency, replacement of a principal in the first year – and certainly in the first weeks – following training (something which happens with surprising frequency) would be reason to cancel the grant.

Summer Reading Institutes

ARI Director Katherine Mitchell is contemplating summer reading institutes, during which she and other senior ARI presenters work with struggling readers both to help the students involved and as a way of offering hands-on instruction to candidates for ARI reading specialist certification. This is an excellent idea and would work well in combination with the proposed ARI reading certification program.

Candidates for ARI reading certification might also spend a week or two at a time as guest teachers at ARI gold star schools. By staying at the school for more than a day or so, they could offer assistance to the school's reading program as well as learning from its success.

Reading Curriculum Materials and Software

The ARI has not required schools to adopt any particular curriculum (such as Voyager or Language!) or any particular set of software (such as Lexia) or reading materials. After extensive discussions with teachers on the question of prepared materials, I believe that this was a wise choice. The very best teachers prefer to prepare their own materials and would resent being told they had to use a particular program.

That said, newer teachers (and perhaps most teachers in less successful schools) would be helped by a good set of instruction materials that embody the lessons taught at the ARI summer training. While the ARI should not require any particular set of instruction materials, the ARI staff, working with panels of teachers from ARI schools, should pre-screen education software and teaching materials and come up with a list of programs that are consistent with the ARI philosophy and that the ARI teacher panels are comfortable with. They'll want to review the material for its level (for example, users of Voyager feel that it is pitched too low) and for its flexibility (some users of Language! liked many of its elements but felt that it was too hard to customize it).

To be helpful to schools, the list of preferred programs should be reasonably short. Several teachers in focus groups objected (strongly) to requiring certain education packages, but they liked the idea of making the selection process easier for schools by having a pre-screening process.

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

More ARI Staff

With a director and 4 education specialists, the ARI central staff is not large enough to work with struggling schools, to select and supervise the recommended 23 in-service center reading specialists, to select the ARI Gold Star schools, to design and run the principal and central officer administrator training programs, to prepare the ARI Reading specialist certification, to run the proposed summer institute, to organize the 3-year renewal training for existing schools, to select new schools, to organize summer training for new schools, and to train and pre-screen ARI presenters. A staff of 8 or 9 is called for.

In addition to this central staff, the ARI should have at least one field staff person for every 20 schools in the initiative. At the current level of 424 schools, this would mean 21 specialists in all, an addition of 10 to the current total. Although they should continue to be located regionally across the state, these specialists should be directly selected, supervised, and employed by the ARI to assure uniform quality and to make them full members of the ARI team.

APPENDIX 1 – QUANTITATIVE ANALYSIS IN MORE DETAIL

Using Standardized Test Scores to Evaluate Individual Schools

The only statewide quantitative information on student performance in Alabama is the Stanford 9 test given to all students in 3rd grade and above. This is a nationally normed test that gives an objective view of how well Alabama students perform. It does not report on absolute performance levels (Susan is a third grader but reads at a 4th grade level); rather, it reports on student performance relative to a comparable group of national students. (Susan scores at the 60th percentile in reading among a national group of 3rd grade students).

If the state's goal is to have every student performing in the 5th stanine or above (which translates to the 45th percentile or higher), the Stanford 9 can measure how close the state or an individual school comes to the goal.

Measuring performance against an absolute standard of this kind, however, is not necessarily an accurate or a fair way to measure the performance of an individual school. Everyone knows that poverty plays an overwhelming role in predicting student performance, but it is still easy to under-estimate just how large this impact is. Looking (as an example) at reading comprehension scores for students at all Alabama schools in 1999, we find using regression analysis that an all-white school with no students in poverty would be expected to have an NCE (normal curve equivalent) score of 60.8; this would translate to a score at the 69th percentile nationally. If this school had all of its children in poverty, it could expect an NCE

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

score of 43.9, which would mean the 39th percentile. If it were all black and all poor, the expected NCE score would fall to 37.2 – the 27th percentile.

If schools and their faculties are judged by their absolute scores alone, they are penalized simply because of their demographics. It is of course appropriate to set a goal that all students can read at or above grade level, as the ARI has done. But it would not be appropriate to judge a faculty's performance toward this goal by absolute scores alone, without taking into consideration school demographics. To take an extreme example, an all-black, all-poor school that scored at the 45th percentile would be somewhat below average if judged by absolute scores alone, and would have about half of its students in the bottom 4 stanines. Yet its students would be scoring almost 20 percentiles higher than the typical school with these demographics. By this measure it should be a success and not a failure.

One way to judge how well school faculties are performing (and, for the purpose of this evaluation, whether ARI schools outperform those not in the program) would be to use regression analysis to predict a school's reading scores based on its demographics and then to compare actual scores with predicted scores.

This would be much fairer (and more reflective of the faculty's efforts) than using absolute scores. Nonetheless, it still has problems when used over time to measure whether a school is improving. The major difficulty is that comparing 3rd grade tests this year with 3rd grade tests last year is comparing the performance of different students. We know that each class of students is unique – one class may have a large number of very bright students; the next may have more students with special needs. Such comparisons are particularly difficult in smaller schools, or at a time when the criteria for including special needs students changes from one year to the next.

Another approach is to compare the progress of individual students over time. This is possible with Alabama's Stanford 9 data, since the test records include the name, social security number, special education status, ethnicity, and school of each student, as well the test results.[7] With this information, it is possible to match individual student test results in any given year with results for the same student in the prior year, and to aggregate this data by grade and by school. This approach is called the "value-added" approach, since it measures how far the school moved the students in its charge. This evaluation uses the value-added approach.

There are several advantages to this method. The comparison is on an apples-to-apples basis; schools are measured on how they improve the scores of the same students over time. The issues of poverty, ethnicity, and special education status are essentially neutralized, since

[7] The test records do not record whether a student was eligible for free and reduced cost lunch, but it may be possible to merge test records with school lunch records.

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

each student is compared with his or her earlier performance. Since the comparison is limited to students who were tested in consecutive years, the scores cannot be biased when the number of special needs students tested is increased from one year to the next or when the demographic composition of the school changes. Finally, since this approach follows individual students, it is possible to see whether school performance is any better for students who attended the school both in the current and prior years (that is, excluding those students who moved into the school in the year preceding the test).

The major drawback of this approach, at least in Alabama, is that there are no Stanford 9 tests given before the 3rd grade. Since the value-added approach measures the change in student performance, the first reading of a student's progress comes in 4th grade. For a K-5 elementary school with equal numbers of students in each grade, this means that the system can shed no light on 2/3 of its students. Worse, if one school does a wonderful job with early literacy so everyone is reading on grade level by 3rd grade, it might show only normal progress in grades 4 and 5. A school with similar demographics that didn't do well with early literacy might well show somewhat better relative performance in grades 4 and 5, even though its students were still below grade level. In effect, the first school is penalized for its success in the (unmeasured) early grades.[8]

Results from the "Value-Added" Approach

The results presented here are based on the value-added approach.[9] As indicated above, the Stanford 9 test results are normalized; the results tell us how Alabama students in a particular grade compared to their counterparts in that same grade all across the country. The test does not measure absolute performance level. We cannot use it to say that the a fourth grade student this year is performing at the level, for example, of 5th grade, second month, and that this is a major improvement over last year, when (at the end of 3rd grade) her performance was, say, 2nd grade, 6th month.

In the example above, the student may be performing at the 55th percentile this year, as against the 45th percentile last year. We can infer that she made an above-average gain because her percentile score increased. A student scoring at the 50th percentile each year is presumably making average progress.

[8] Measuring school performance by student scores relative to regression predictions, particularly for 3rd grade students, would shed some light on primary year success, since schools that did a good job in the early years would presumably show higher 3rd grade scores. However, turnover is so high that a significant percentage of students who had been in the school in 1st grade would no longer be there in 3rd grade.

[9] The Department of Education has requested additional analysis of the Stanford 9 and early literacy data using a variety of approaches. It has also asked that the report be presented in its current form to allow work to begin on implementing the evaluation's recommendations. The additional statistical analysis will be available by March of 2002.

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

In fact, the data are somewhat more complicated. In some years, Alabama students average slightly above the 50th percentile; in other years, slightly below. If a student jumps from the 45th to 55th percentile in one year, this may mark major progress; it may also reflect the fact that Alabama students generally did poorly the first year and better the second.

In this analysis, the performance of students in any particular grade level in Alabama schools not participating in the ARI is taken as a benchmark; performance of ARI schools, both collectively and individually, is measured against this benchmark. For example, 6th graders in 2001 in non-ARI schools (for whom we also had test records a year earlier) had a normal-curve equivalent in reading comprehension of 52.1.[10] In 2000, as 5th graders, these same students had scored 51.0. For our analysis, this gain of 1.1 from 5th to 6th grades was taken as the norm; a school with this score was assigned a zero. If an ARI school showed a gain of 2.1, then it's score was taken as 1.0, meaning its students gained 1 point on the NCE relative to non-participating schools.

The results using this approach are shown in the table below. Results are combined across years but are in fact calculated separately for each year. Within each year, scores are first calculated by grade and compared with the norm (that is, non-ARI schools) for that year and grade. The results are then aggregated across grades to produce school totals. The table below shows the average for all ARI schools, by cohort and then by grade grouping (elementary, middle, high). The table also shows the performance individually of the cohort 1 and 2 schools. Individual school names are not disclosed; the individual school data is included to give a sense of the differences across schools.

In 1999, the 14 Cohort 1 schools had a score of .1, meaning their results that year were all but identical to those of non-ARI schools. The table also shows us that there were 2,718 students tested at those 14 schools in that year for whom test records are available from the prior year. Generally, 3rd graders would be excluded from this total (unless they had failed and repeated 3rd grade); so are students who had not been in Alabama a year earlier. At the ARI schools in cohort 1, 49% of the students are poor and 47% are black. xxx

In the second year, students at both cohort 1 and cohort 2 outperformed the norm (that is, scored better than students at non-ARI schools for whom data is available for both years). Performance in this second year was particularly strong in the cohort 1 schools.

---

[10] The normal curve equivalent is based on the percentile results, but is designed to reflect roughly equal increments in student performance. Since most student scores are near the mean, a relatively small change in performance can make a major change in the student's ranking. A change of 1 NCE point translates to a gain of 1.77 percentiles in the middle of the distribution. With very high or very low scores, however, even a big difference in performance may not make much of a dent in percentile rankings. A student with an NCE score of 45 is in fact performing just above the 41st percentile.

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

In the third year, almost 49,000 students in ARI schools (for whom we have prior-year results) took the test. Taken as a group, the 14 cohort 1 schools actually lost ground, while the cohort 2 schools did slightly better than non-participants, and the cohort 3 schools outperformed the non-participants by half an NCE point.

The table also shows cumulative gains; by the third year, cohorts 1 and 2 were ahead of non-participants by 1.1 points. Since the number of years any school has been in the ARI varies, a useful way to compare performance is the average annual gain (or loss) relative to the norm. On average, ARI schools have risen student scores relative to non-ARI schools by .5 NCE points per year.[11] For students near the mean, this translates to a gain of about 1 percentile. Put another way, if the average student at a non-participating school is in the 50th percentile each year as she moves up from grade 3, the average student at an ARI school will be in the 51st percentile after one year in the program, the 52nd percentile after 2 years, and the 53rd percentile after 3 years.

The data suggest that the gains at any ARI school will level off over time. The cohort 1 schools lost ground in year 3; cohort 2 gains were less in year 2 than year 1. The first year gain of the cohort 3 schools was less than that for the cohort 2 schools a year earlier. Given this pattern, it is unlikely that a student who starts at an ARI school at the 50th percentile (NCE Score 50) in Grade 3 and stays at ARI schools throughout his career will gain 1 percentile every 2 years, moving to an NCE score of 54 in 11th grade (a percentile score of 57).

[11] This average is weighted to reflect the increasing number of ARI students in successive years.

CAPE ANN ECONOMICS

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

		98-99	99-00	00-01	Cumulative 98-99	Cumulative 99-00	Cumulative 00-01	Wtd Avg Yrly Ch	% of Students Black	% of Students Poor	Students Tested 98-99	Students Tested 99-00	Students Tested 00-01
All ARI Schools		**0.1**	**0.8**	**0.4**				**0.5**	**42%**	**50%**	**2,718**	**15,767**	**48,824**
Cohort 1		0.1	1.6	-0.7	0.1	1.8	1.1	0.4	49%	47%	2,718	2,114	2,270
Cohort 2			0.8	0.3		0.8	1.1	0.5	37%	44%		13,653	13,348
Cohort 3				0.5			0.5	0.5	43%	52%			33,206
Elementary			0.1	0.8	0.2			0.3	43%	54%	1,733	7,477	26,147
Middle			0.1	0.8	0.7			0.7	41%	45%	751	6,382	13,647
High			0.4	1.1	0.4			0.5	30%	30%	234	1,908	9,030
ARI Cohort 1													
Cohort 1 Sch 1	Elementary		5.2	1.9	1.2	5.2	7.1	8.3	2.8	19%	57%	115	107	104
Cohort 1 Sch 2	Elementary		0.6	4.2	2.8	0.6	4.9	7.7	2.6	48%	43%	109	110	95
Cohort 1 Sch 3	Elementary		1.7	1.1	2.3	1.7	2.8	5.0	1.7	91%	79%	55	72	57
Cohort 1 Sch 4	High		-0.2	4.4	-1.3	-0.2	4.2	2.9	1.0	99%	80%	337	368	346
Cohort 1 Sch 5	Elementary		0.0	1.9	-0.1	0.0	1.9	1.8	0.6	52%	38%	103	96	93
Cohort 1 Sch 6	Elementary		0.1	0.5	0.5	0.1	0.7	1.2	0.4	41%	49%	203	175	177
Cohort 1 Sch 7	Elementary		1.1	-2.3	1.8	1.1	-1.2	0.6	0.2	29%	52%	270	127	131
Cohort 1 Sch 8	Middle		-0.7	1.7	-0.8	-0.7	0.9	0.2	0.1	14%	7%	704	543	745
Cohort 1 Sch 9	Elementary		-0.2			-0.2	-0.2	-0.2	-0.2	17%	37%	362	1	2
Cohort 1 Sch 10	Elementary		-1.3	4.3	-3.3	-1.3	3.0	-0.3	-0.1	47%	36%	110	132	126
Cohort 1 Sch 11	Elementary		-1.8	1.7	-0.8	-1.8	-0.1	-0.9	-0.3	73%	76%	73	87	91
Cohort 1 Sch 12	Elementary		-0.4	-1.3	0.8	-0.4	-1.7	-0.9	-0.3	100%	40%	135	136	117
Cohort 1 Sch 13	Elementary		3.3	-3.0	-1.6	3.3	0.3	-1.3	-0.4	98%	94%	75	83	95
Cohort 1 Sch 14	Elementary		0.0	-1.9	-10.1	0.0	-1.9	-12.0	-4.0	99%	107%	67	77	91

		Change			Cumulative			Avg Ch	% of Students		Students Tested		
		98-99	99-00	00-01	98-99	99-00	00-01	Per Yr	Black	Poor	98-99	99-00	00-01
ARI Cohort 2													
Cohort 2 Sch 1	Elementary		8.6	3.9		8.6	12.5	6.2	100%	48%		104	126
Cohort 2 Sch 2	Elementary		4.7	2.9		4.7	7.5	3.8	63%	48%		55	49
Cohort 2 Sch 3	Elementary		2.4	3.0		2.4	5.5	2.7	51%	85%		92	97
Cohort 2 Sch 4	Elementary		2.2	2.3		2.2	4.5	2.3	10%	67%		139	118
Cohort 2 Sch 5	Elementary		2.8	1.5		2.8	4.3	2.1	43%	59%		98	92
Cohort 2 Sch 6	Elementary		3.4	0.8		3.4	4.1	2.1	100%	96%		143	139
Cohort 2 Sch 7	Elementary		0.4	3.4		0.4	3.8	1.9	63%	75%		79	92
Cohort 2 Sch 8	Middle		3.1	0.6		3.1	3.7	1.9	35%	39%		581	589
Cohort 2 Sch 9	Elementary		2.3	1.3		2.3	3.7	1.8	5%	57%		88	79
Cohort 2 Sch 10	Middle		1.0	2.6		1.0	3.6	1.8	62%	59%		421	307
Cohort 2 Sch 11	High		1.9	1.6		1.9	3.5	1.8	0%	52%		211	192
Cohort 2 Sch 12	Elementary		2.5	1.0		2.5	3.5	1.7	0%	20%		257	264
Cohort 2 Sch 13	Elementary		3.1	0.3		3.1	3.3	1.7	24%	64%		132	121
Cohort 2 Sch 14	Elementary		1.6	1.5		1.6	3.2	1.6	99%	100%		65	46
Cohort 2 Sch 15	Middle		1.1	2.0		1.1	3.0	1.5	51%	64%		310	335
Cohort 2 Sch 16	Elementary		3.3	-0.5		3.3	2.8	1.4	96%	86%		54	40
Cohort 2 Sch 17	Elementary		2.0	0.7		2.0	2.7	1.4	6%	43%		104	112
Cohort 2 Sch 18	Elementary		3.4	-0.7		3.4	2.7	1.4	22%	57%		109	110
Cohort 2 Sch 19	Middle		1.8	0.8		1.8	2.6	1.3	27%	58%		379	408
Cohort 2 Sch 20	Elementary		1.8	0.7		1.8	2.4	1.2	64%	41%		152	150
Cohort 2 Sch 21	Elementary		1.2	1.2		1.2	2.4	1.2	23%	48%		116	101
Cohort 2 Sch 22	Elementary		4.5	-2.2		4.5	2.3	1.1	49%	76%		61	71
Cohort 2 Sch 23	Elementary		1.3	0.9		1.3	2.2	1.1	37%	68%		232	214
Cohort 2 Sch 24	Elementary		1.3	0.8		1.3	2.1	1.1	14%	19%		117	107
Cohort 2 Sch 25	Elementary		1.4	0.7		1.4	2.1	1.0	59%	92%		54	69
Cohort 2 Sch 26	Elementary		1.2	0.8		1.2	2.0	1.0	41%	80%		77	84
Cohort 2 Sch 27	High		0.0	1.9		0.0	1.9	1.0	0%	38%		176	168

					Cumulative			Wtd Avg	% of Students		Students Tested		
		98-99	99-00	00-01	98-99	99-00	00-01	Yrly Ch	Black	Poor	98-99	99-00	00-01
Cohort 2 Sch 28	Elementary		2.3	-0.5		2.3	1.8	0.9	31%	53%		186	197
Cohort 2 Sch 29	Elementary		1.0	0.7		1.0	1.7	0.9	100%	43%		142	132
Cohort 2 Sch 30	High		2.3	-0.7		2.3	1.6	0.8	13%	42%		109	105
Cohort 2 Sch 31	High		0.7	0.8		0.7	1.5	0.7	17%	16%		524	535
Cohort 2 Sch 32	High		0.9	0.4		0.9	1.3	0.6	42%	43%		359	368
Cohort 2 Sch 33	Elementary		3.9	-2.7		3.9	1.2	0.6	29%	34%		56	63
Cohort 2 Sch 34	Middle		0.5			0.5	0.5	0.5	12%	10%		859	836
Cohort 2 Sch 35	Elementary		0.7	0.3		0.7	1.0	0.5	36%	39%		126	140
Cohort 2 Sch 36	Middle		1.3	-0.4		1.3	1.0	0.5	21%	20%		663	657
Cohort 2 Sch 37	Elementary		-0.8	1.8		-0.8	1.0	0.5	97%	65%		233	227
Cohort 2 Sch 38	Middle		-0.3	1.1		-0.3	0.8	0.4	27%	26%		376	361
Cohort 2 Sch 39	Elementary		-0.1	0.9		-0.1	0.8	0.4	6%	3%		564	546
Cohort 2 Sch 40	Elementary		-1.3	2.1		-1.3	0.8	0.4	57%	38%		249	231
Cohort 2 Sch 41	Elementary		1.4	-0.8		1.4	0.6	0.3	3%	8%		189	187
Cohort 2 Sch 42	Elementary		1.8	-1.2		1.8	0.6	0.3	82%	89%		132	132
Cohort 2 Sch 43	Elementary		0.6			0.6	0.6	0.6				129	0
Cohort 2 Sch 44	Elementary		-1.8	2.3		-1.8	0.5	0.3	3%	47%		76	83
Cohort 2 Sch 45	Elementary		0.5	-0.3		0.5	0.2	0.1	29%	30%		333	327
Cohort 2 Sch 46	Elementary		1.1	-1.0		1.1	0.1	0.1	74%	82%		124	108
Cohort 2 Sch 47	Elementary		3.2	-3.2		3.2	0.0	0.0	100%	65%		88	102
Cohort 2 Sch 48	Elementary		-1.1	0.9		-1.1	-0.2	-0.1	20%	68%		280	271
Cohort 2 Sch 49	Elementary		1.1	-1.5		1.1	-0.3	-0.2	99%	87%		90	71
Cohort 2 Sch 50	Elementary		-0.1	-0.2		-0.1	-0.4	-0.2	1%	40%		285	329
Cohort 2 Sch 51	High		-0.2	-0.2		-0.2	-0.4	-0.2	8%	12%		657	668
Cohort 2 Sch 52	Elementary		0.1	-1.5		0.1	-1.4	-0.7	14%	54%		142	152
Cohort 2 Sch 53	Elementary		4.2	-5.9		4.2	-1.6	-0.8	52%	78%		70	54
Cohort 2 Sch 54	Middle		-1.0	-1.1		-1.0	-2.1	-1.0	29%	11%		617	602
Cohort 2 Sch 55	Elementary		0.3	-2.9		0.3	-2.6	-1.3	93%	92%		97	95
Cohort 2 Sch 56	Middle		-1.2	-1.5		-1.2	-2.7	-1.4	45%	58%		501	503
Cohort 2 Sch 57	Middle		-1.9	-1.4		-1.9	-3.3	-1.6	61%	66%		851	826
Cohort 2 Sch 58	Elementary		-6.5	2.9		-6.5	-3.6	-1.8	90%	84%		75	72
Cohort 2 Sch 59	Elementary		0.2	-6.2		0.2	-6.0	-3.0	36%	65%		62	88

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

The table also gives basic demographic data for the ARI schools. All together, they had a 50% poverty rate; 42% of their students (who took the test) were black. Comparable numbers for non-ARI schools are 49% poverty, 32% black.

The first table also summarizes the results by grade level. ARI schools actually do better relative to non-participants in middle grades and high school than they do in elementary schools.

The data for schools in cohorts 1 and 2 is sorted by descending order of cumulative performance. The best 2 schools in Cohort 1 had moved their students ahead of the norm by 8 NCE points after 3 years (a gain of 14 percentiles). Neither of these schools has a low poverty rate; both are integrated. The lowest performing ARI school made no progress the first year and had a disastrous year in 2000-2001; its students have lost 12 NCE points (21 percentiles) in 2 years.

The very best cohort 2 school, like the worst cohort 1 school, is all black, although its poverty rate is "only" 50%. It has moved its students ahead by 12.5 NCE points in just 2 years. Of 59 cohort 2 schools, only 12 scored below the norm set by non-ARI schools.

More Detailed Analysis

A more detailed analysis has been prepared, showing actual scores of ARI and non-ARI schools by grade level by year. A sample of this analysis, showing grades 4 and 5 for the year 2000-2001, is displayed below.[12]

Reading across, there were 1,036 students in cohort 1 schools who took the test in the spring of 2001 as 4th or 5th graders and whose scores can be matched with test results a year earlier. Of these, 57% are poor and 62% black – much higher than the 49% of students in poverty and 32% of students who are black in non-participating schools. Their scores rose from an average of 49.6 last year to 49.9 this spring – a gain of .3, which is slightly less than the gain of .4 in non-ARI schools. These schools lost ground this year relative to non-ARI schools, as we've seen in the data reported earlier.

[12] Space precludes a full exposition of the data available. This complete table covers all Alabama schools with 4th and 5th grades (including those not in the ARI); this table exists also for the 2 prior years. A similar set of 3 tables has been produced for Grade 6, for Grades 7 and 8, and for grades 9, 10, and 11.

CAPE ANN ECONOMICS

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

ARI Comparison 4th and 5th graders in 00-01 who also were tested in 99-00 - ARI vs Other Schools

	Students			Average Reading Comprehension						Studs < 40th Pcntile, 00			Deficit from 40th Pcntile			
				All Students			Studs in Same Sch		Black Studs	% Tot	Avg RC	% Inc 10	Pct			
	Number	Pct Black	Pct Poor	Ch	01	00	% Tot	Ch	Ch	Studs	Ch	or more	Ch	Ch	01	00
Cohort 1	1,036	62%	57%	**0.3**	49.9	49.6	87%	0.0	-0.4	43%	5.1	31%	-0.2	-3%	5.3	5.5
Cohort 2	6,181	38%	49%	**0.6**	52.3	51.7	82%	0.6	0.3	39%	4.3	27%	0.0	0%	5.0	4.9
Cohort 3	18,595	41%	54%	**0.8**	51.2	50.4	78%	0.9	0.9	42%	5.0	31%	-0.3	-6%	5.4	5.7
Not in ARI	**67,941**	**32%**	**49%**	**0.4**	**51.7**	**51.3**	**78%**	**0.6**	**0.1**	**39%**	**4.9**	**31%**	**-0.1**	**-1%**	**5.1**	**5.2**
Difference from Baseline (not in ARI)																
Cohort 1		30%	8%	**-0.1**	-1.8	-1.6	9%	-0.6	-0.6	4%	0.2	1%	-0.1	-2%	0.1	0.3
Cohort 2		6%	0%	**0.2**	0.7	0.5	5%	-0.1	0.2	-1%	-0.6	-3%	0.1	2%	-0.2	-0.3
Cohort 3		9%	4%	**0.4**	-0.5	-0.9	1%	0.2	0.7	2%	0.1	0%	-0.3	-5%	0.2	0.5
ARI Cohort 1																
Cohort 1 Sch 1	93	53%	43%	**5.1**	52.8	47.6	96%	5.4	5.3	47%	8.1	39%	-2.6	-43%	3.5	6.1
Cohort 1 Sch 2	131	29%	52%	**4.3**	51.7	47.4	93%	3.7	5.8	49%	9.8	36%	-2.3	-35%	4.3	6.6
Cohort 1 Sch 3	55	95%	79%	**3.0**	42.4	39.4	78%	3.2	3.8	71%	4.4	26%	-1.4	-16%	7.1	8.5
Cohort 1 Sch 4	113	49%	49%	**2.9**	53.7	50.8	82%	3.5	4.2	38%	8.3	49%	-0.9	-20%	3.8	4.7
Cohort 1 Sch 5	68	13%	57%	**2.3**	55.0	52.7	84%	2.0	-3.7	34%	10.0	48%	-1.5	-34%	2.9	4.5
Cohort 1 Sch 6	115	99%	40%	**1.1**	42.1	41.0	83%	1.5	1.1	63%	3.8	26%	0.0	0%	8.7	8.8
Cohort 1 Sch 7	92	53%	38%	**0.1**	50.9	50.7	85%	-0.7	-0.3	40%	2.3	35%	1.2	39%	4.4	3.2
Cohort 1 Sch 8	89	72%	76%	**-0.1**	47.1	47.1	83%	-0.9	1.4	52%	3.1	20%	-0.7	-10%	6.8	7.5
Cohort 1 Sch 9	126	47%	36%	**-2.9**	61.5	64.4	94%	-3.2	-3.1	17%	0.7	14%	0.5	34%	2.1	1.6
Cohort 1 Sch 10	65	100%	94%	**-4.1**	42.2	46.4	94%	-4.3	-4.1	49%	1.2	28%	2.0	33%	8.0	6.0
Cohort 1 Sch 11	89	99%	107%	**-10.1**	41.4	51.6	81%	-12.8	-10.2	29%	-1.2	15%	4.4	107%	8.5	4.1

49 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

Of the 1,036 students in the cohort 1 schools, 87% had been in the same school a year earlier; these same-school students showed no gain in score. The students who had been in 3rd grade in the same school a year earlier actually did slightly *worse* than students who had moved in since the year-earlier test. For the other ARI cohorts, and for non-ARI schools, students who'd been in the same school the prior year did slightly (but only slightly) better than the recent arrivals. Because these differences are so small, the study has not made this distinction in the general conclusions presented earlier.

The turnover rate for non-ARI schools and for schools in cohort 3 was 22% - only 78% of students who'd taken the test somewhere in Alabama the year before had done so at the same school. A turnover rate consistently this high over time would mean that in 3 years fewer than half the original students remained. This is why it is difficult to make inferences about how well schools are doing in 1st and 2nd grade by looking at their 3rd grade scores a year or two later.

The black students in these cohort 1 schools showed a drop of .1 while the school as a whole showed a gain of .3. There's a similar pattern (black gains less than white gains) in cohort 2 schools and in non-ARI schools, while cohort 3 schools showed the greatest gain, .8 NCE points for all students and .9 NCE points for black students. ARI cohort 2 and cohort 3 schools on average showed greater progress, for black students and for all students, than did non-ARI schools.

This more detailed analysis allows us to look at how ARI schools are doing with their struggling readers. Looking only at students who had been below the 40th percentile in the spring of 2000 (43% of the students in cohort 1 schools; about the same for other cohorts and for non-ARI schools), the gain in scores was 5.1 for cohort 1 schools – much higher than the .3 gain for all students in these schools. However, this higher gain does not distinguish ARI from non-ARI schools, which showed a gain of 4.9 for their struggling readers. The picture is no different for cohort 2 and cohort 3 schools. A look farther down the table, however, shows that the highest performing cohort 1 schools did very well with their struggling readers.[13]

Another way to look at performance with struggling readers is to see how many struggling readers made very large year-to-year gains. Thirty-one percent of the struggling readers at the cohort 1 schools made gains of 10 or more NCE points. The results along these lines were essentially the same for all 3 ARI cohorts and for non-ARI schools.

A third approach is to look at how well schools did in eliminating the gap between the lowest acceptable score and actual performance of struggling readers. The school's deficit for

[13] The highest performing cohort 1 schools on this table are not necessarily the same as those in the earlier table, since this table sorts only by performance on the 2001 Stanford 9 tests while the earlier table ranks cohort 1 schools by their 3-year cumulative performance.

40 River Road
Annisquam Village
Gloucester, MA 01930
(978) 281-5004

struggling readers is defined as the difference between a score at the 40th percentile (which is NCE 45) and the student's actual score, divided by the total number of students in the school. Thus, the deficits of about 5 percentage points occur because about 40% of students score below the 40th percentile and have an average shortfall of about 14 points.

In this instance, a negative score indicates success – the school succeeded in reducing the extent to which students score below the 40th percentile. Except for cohort 3 schools, there is no significant difference between ARI and non-ARI schools in decreasing this deficit.

The scores of some individual schools are very impressive. Three of the cohort 1 schools reduced the low-performance deficit by 33% or more; there was essentially no deficit reduction in non-ARI schools.